2020

Debating How to Measure Media Exposure in Surveys

Seth K. Goldman
University of Massachusetts - Amherst

Stephen Warren

Debating How to Measure Media Exposure in Surveys

Seth K. Goldman and Stephen M. Warren

Abstract

To answer many of the most pressing questions in the social sciences, researchers need reliable and valid measures of media exposure that can be implemented in surveys. Despite considerable effort, however, substantial disagreement remains about how best to measure this key concept. This chapter critically reviews the debate surrounding traditional frequency measures of exposure to "news" and contemporary list-based measures of political media exposure. It also evaluates the related debate over how best to capture the effects of media exposure with different observational research designs. Overall, the chapter finds that although substantial progress has been made in measurement and research design, both issues require more attention if scholars are to understand the many and varied effects of media exposure.

Keywords

media exposure; surveys; research design; media effects; news media; campaigns

Media exposure is at the core of many of the most pressing questions in the study of politics and society. For instance, do media promote tolerance or intolerance toward marginalized social groups? Has an increasingly partisan media environment reduced public exposure to a diversity of viewpoints? Are media at least partly to blame for political polarization? Is the high-choice media environment encouraging political engagement, apathy, or cynicism? Does media exposure to impersonal others lead to political trust or wariness? Do coarse and uncivil media make it harder for people to empathize with the other side? How does election coverage influence citizens' views of leaders and democratic institutions? Answering these and other questions requires reliable and valid measures of media exposure, yet the meaning and measurement of this concept remains a matter of debate (de Vreese and Neijens 2016; Jerit and Barabas 2011).

In this chapter, we critically assess several measures of media exposure to politics (for a review of exposure measures across topical domains, see Nagler 2018). Our focus is on measures that can be implemented in population-based surveys, which are necessary for testing questions about media effects in real-world settings. We evaluate evidence from research that tests the reliability or validity of one or more measures of political media exposure. Finally, we discuss the relative merits of different survey-based research designs for examining the effects of media exposure on attitudes and behaviors. Our review reveals significant progress in the measurement of political media exposure, as well as in the design of observational media effects studies. However, few studies focus on these methodological issues—a serious problem that we hope this chapter encourages scholars to address more deeply in the future.

Traditional Measures

The conventional approach to measuring media exposure relies on traditional frequency measures in which survey respondents are asked how many hours per day or days per week they watch, read, or listen to the "news" on a given medium. The most recent iteration of this approach, developed by Althaus and Tewksbury (2007a, 2007b) and included on the 2012 American National Election Study (ANES), asked respondents in a series of questions how many days in a typical week they watch, read, or listen to "news on the internet," "national news on TV," "news in a printed newspaper," and "news on the radio." Although the 2016 ANES dropped these items due to longstanding concerns about their reliability and validity, many scholars continue to use these and similar measures to study the effects of media exposure (e.g., Edgerly et al. 2017; Kruikemeier and Shehata 2017).

The reliability of traditional exposure measures has been in doubt since Bartels (1993) identified measurement error as a key explanation for the limited evidence of media effects in social science research. Notably, Bartels (1993) offers the only analysis of traditional exposure measures that leverages three waves of panel data to calculate "true-score" reliability estimates (Wiley and Wiley 1970), which are unique in their ability to separate measurement error from true changes in individual levels of media exposure over time. Other types of reliability estimates either cannot distinguish between inconsistency in responses over time and real change (e.g., test-retest reliability) or only gauge the correlation between different survey items at a single point in time (e.g., internal consistency). True-score reliability is thus a much higher standard because it estimates the extent to which a measure captures the true values of the variable (Zaller 2002). Bartels (1993) reported a true-score reliability of 0.75 for the 1980 ANES measure that asked, "How often do you watch the national network news on early evening TV—every evening, 3 or 4 times a week, once or twice a week, or less often?"

Scholars often cite Bartels (1993) as evidence of the unreliability of traditional exposure measures, but the true-score reliability of the TV exposure measure actually compared favorably to several other measures of political constructs. For example, Bartels (1993) found lower levels of true-score reliability for standard measures of presidential job approval and issue preferences than for TV exposure. Nonetheless, once Bartels (1993) corrected for measurement error in the TV exposure measure, he found much stronger evidence of media effects on political opinions.

The unreliability of traditional exposure measures may be overblown, but so too are claims of high reliability based on assessments of internal consistency using Cronbach's alpha, which relies on the strength of inter-item correlations. Using this metric, scholars often report high reliabilities for scales of traditional exposure items. For example, Zaller (2002) found that the four 1996 ANES days per week items had an alpha reliability of 0.79, and other studies have reported similar estimates from multi-item scales (e.g., Chaffee and Schleuder 1986; Price and Zaller 1993). However, assessments of internal consistency only show whether several items hang together, not whether people are able to consistently report their exposure over time, as true-score reliability does. In any case, as the meaning of "news" has become widely debated, peoples' ability to reliably report their levels of news exposure has likely deteriorated.

Reliability aside, the major focus of criticism of traditional exposure measures is their validity, of which the gold standard is political knowledge *gain* under the assumption that exposure to news should impart political information. Traditional exposure measures—even those with purportedly high reliability—typically show only weak correlations with political knowledge (Zaller 2002). One well-known study examined whether traditional exposure items predict recall of current events that had been featured in news coverage (Price and Zaller 1993). Across many analyses, the frequency of newspaper and national TV news exposure produced small and rarely significant effects on news recall, leading the authors to conclude that scholars should abandon the use of self-reports of media exposure altogether (Price and Zaller 1993; see also Chang and Krosnick 2003).

Interestingly, where many see evidence of weak and inconsistent effects, others see evidence of reasonably modest effects. Althaus and Tewksbury (2007b, 8) interpreted Price and Zaller's (1993) analyses as demonstrating that "media use adequately predicts knowledge in and of itself." Yet absent a theory as to why traditional exposure measures are only inconsistently related to news recall, it is not evident why Price and Zaller's (1993) findings should be reinterpreted as evidence of the measures' validity. Althaus and Tewksbury (2007a, 2007b) also reported analyses of the 2006 ANES pilot study that showed some significant relationships between traditional exposure measures and perceptions of issue differences between candidates, interpersonal political discussion, and political participation, Yet many of the relationships were not significant, providing ambiguous evidence of validity overall.

Although other cross-sectional analyses have found significant relationships between traditional exposure measures and political knowledge (Chaffee and Schleuder 1986; Eveland et al. 2005; Eveland, Hutchens, and Shen 2009; Tewksbury, Althaus, and Hibbing 2011), the implications of these analyses are unclear because they provide weak causal evidence. As Bartels (1993, 267) emphasized, "differences in opinions between those exposed to the media and those who remain unexposed may simply reflect preexisting differences between the two groups in political attitudes or characteristics." Even after including control variables, unmeasured and/or unobservable variables may still produce spurious associations. Reverse causality is also a major concern given that those with more preexisting knowledge are more likely to seek out and retain new information. The few panel studies that have been employed rely on lagged-dependent variable models (LDVMs) which do not overcome these problems either (Allison 1990, 2000). Moreover, LDVMs only capture whether the rank-order of individuals changed over time, not whether the same individuals exhibited knowledge gains. In sum, prior research has not convincingly demonstrated that traditional exposure measures cause political knowledge gain, leaving the predictive validity of these measures in question.

Traditional exposure measures have also been criticized on the grounds of low convergent validity due to overreporting. Price and Zaller (1993) reported that 35 percent of respondents in the 1989 ANES pilot study said they listened to National Public Radio (NPR) versus 6 percent according to Arbitron's weekly diary data. They also found overreports of reading the *Wall Street Journal* compared to the newspaper's internal estimates. For national TV news, Prior (2009a) compared ratings from Nielsen's people-meter data to a question from the 2000 National Annenberg Election Survey (NAES), which asked "How many days in the past week did you watch the national network news on TV—by national news, I mean Peter Jennings on ABC, Dan Rather on CBS, Tom Brokaw on NBC, Fox News or UPN News?" Based on Nielsen's ratings, Prior (2009a) estimated that on an average weekday between 30 and 35 million people watched one of the three broadcast news programs—UPN and the Fox broadcast network did not have such programs—whereas between 85 and 110 million did according to the NAES. These comparisons are problematic, however, because surveys and ratings data use different metrics and often refer to different sources; moreover, the ratings data themselves have been criticized as unreliable and lacking validity (*Economist* 2013; Milavsky 1992; Napoli 2003).

Still, some degree of overreporting is likely, and scholars have pointed to two possible explanations: social desirability bias and the high cognitive burden placed on respondents. According to the social desirability hypothesis, respondents inflate their reported news viewing because they perceive a norm of democratic citizenship that includes being knowledgeable of current events. Plausible as this explanation may be, the available evidence does not support it. One survey showed only a weak correlation between an individual's propensity to engage in socially desirable responding and their self-reported news use (Eveland et al. 2009), and a list experiment revealed no differences in self-reported exposure between those asked about their viewing directly versus indirectly (Prior 2009b).

The more likely explanation for overreporting is the high cognitive burden placed on respondents. Answering a traditional exposure question involves a multistage process in which respondents must understand the question—including what counts as “news”—accurately recall their frequency of past exposure, and then map their recollections onto the response options in the question (Schwarz and Oyserman 2001). Experiments embedded in national surveys reveal that respondents do not misreport their exposure for lack of effort but instead because of flawed estimation strategies. People appear to infer from their interest in politics and the (mistaken) assumption that most people watch the news that they probably watch the news, too (Prior 2009b). Giving people more time to think about how often they watch the news did not reduce overreporting, but telling them that most people do not watch the news did.

Importantly, even if people do significantly overreport their news exposure, it may not undercut media effects research insofar as self-report measures still appear to capture more signal than noise. One study compared passive tracking data collected by individuals’ cell phones to a survey question that asked, “How much time do you spend watching TV news in a typical day” (never, fewer than two hours, between two and four hours, or more than four hours) (LaCour and Vavreck 2014). About four-in-ten participants overreported their news exposure, yet most provided either accurate reports or underreports, and self-reports were positively and significantly related to actual exposure. Overall, those who reported higher exposure *did* have a greater likelihood of actually seeing or hearing more news.

Despite this positive result, concerns about the reliability and validity of traditional frequency measures of media exposure remain widespread. The growing fragmentation of the media environment presents an even greater challenge. Measures of exposure to “national TV news” or “news on the internet” are useful only insofar as there are consistent messages across sources within a given medium. Even then, respondents are unlikely to agree on what “news” means anymore, adding considerable uncertainty to the interpretation of any observed effects of exposure. What most scholars ultimately care about is the impact of exposure to particular messages or types of content (Slater 2004), and in the current information environment the content people encounter varies widely across sources. Because traditional measures do not capture exposure to those distinctive messages, we turn to a new class of measures that do.

List-Based Measures

In recent years, scholars have increasingly relied on list-based measures of media exposure in which survey respondents are presented with lists of individual sources, rather than researcher-defined categories, and asked which they use regularly. In one formulation, first included on the 2008 NAES internet panel survey, respondents are provided with lists of the most watched politically relevant television programs—including nightly network news, morning shows, opinion programs, newsmagazines, daytime talk shows, and satire—and asked to check off the ones that they watch regularly, defined as at least once per month (Dilliplane, Goldman, and Mutz 2013). This approach has been extended beyond television to other media, though debate remains about its conceptualization and implementation (Anderson, de Vreese, and Albaek 2016; Guess 2015; Prior 2013).

The benefits of list-based measures are two-fold. First, they capture exposure to a wider range of politically relevant content than traditional exposure measures, while reducing the cognitive burden placed on respondents. Respondents need not understand what counts under the umbrella of news, a category that may have once been obvious in its meaning but which now is highly disputed. Respondents with differing definitions of news need only recognize and check off individual sources that they use regularly. This also alleviates the need for respondents to engage in the mental arithmetic required to calculate frequency of use in days per week or hours per day. As an approach designed to test theories of media effects, regular use captures the type of cumulative exposure most likely to influence attitudes and behaviors.

Second, the list-based approach encompasses the varied ways in which people now encounter political content in the media. In addition to national network news programs, citizens are now exposed to political content through morning shows, back-to-back news programs on cable networks, and even daytime talk shows, which periodically include political issues and candidates, especially during election campaigns. Many people have also turned to alternative formats such as opinion programs and late-night satire that do not claim to be news but which still impart political information. Given the large and ever-changing menu of options available, a measure that allows researchers to add or subtract sources to most appropriately test relevant theories of influence is of increasing value. Additionally, the list-based approach easily captures exposure across the growing number of media platforms, as well as both concurrent and time-delayed media use.

To assess the reliability of the list-based measure of political television exposure, Dilliplane et al. (2013) used three waves of nationally representative panel data fielded over the internet during the 2008 presidential election, gathered as part of the 2008 NAES. They calculated true-score reliabilities (Heise 1969) of 0.83 for an indicator of the total number of politically relevant programs viewed, 0.84 for an indicator weighted by each program's level of campaign content, and 0.88 for the individual programs (averaged across all 49 programs asked about in the survey). Notably, these true-score reliabilities are higher than Bartels (1993) reported for a traditional "days per week" measure of television news exposure (0.75).

For tests of predictive validity, Dilliplane et al. (2013) employed fixed effects models of within-person change (Allison 2009), which showed that within-person increases in the number of programs viewed significantly predicted within-person *gains* over time in knowledge of candidate issue positions. In other words, the same exact individuals who increased in their political television viewing also increased in their levels of political knowledge. Stable levels of exposure based on average viewing across all three waves also predicted knowledge gains. As expected, the measure weighted by level of campaign content produced the strongest results. The number of programs viewed uniquely predicted levels of *visual* candidate knowledge as well (i.e., recognition of candidate faces). In sum, across numerous analyses, the list-based measure demonstrated strong true-score reliability, predictive validity, and discriminant validity.

Debating the Merits of the List-Based Approach

In 2012, the ANES added list-based measures of exposure to political television programs, radio programs, websites, and print and online newspapers to its pre-election survey, in addition to the traditional frequency measures of news exposure. Soon after, Prior (2013) offered a variety of critiques of the list-based approach, and Goldman, Mutz, and Dilliplane (2013) responded in turn.

Prior's (2013) first critique concerns construct validity, or the extent to which a measure taps the underlying concept of interest. He argued that the construct validity of the new measure is low because it places a high cognitive burden on respondents, who are unlikely to remember the names of most programs that they view, and fails to capture frequency of exposure. Because the measure neglects the duration of viewing, someone who watches two programs on one day could be scored as having greater exposure than someone who watches one program every day.

Goldman et al. (2013) responded that the list-based approach was originally designed to measure media exposure among children (e.g., Huesmann et al. 2003) and requires less cognitive effort than traditional measures. The measure's high true-score reliability also suggests that people can, in fact, recognize the programs that they watch regularly. Moreover, two studies that do not rely on self-reports found a close correspondence between the number of programs viewed and the total frequency of viewing. According to an analysis of Dutch people-meter data, the "duration and the number of programs [is] nearly perfectly related" (Wonneberger, Schoenbach, and Meurs 2013, 95; see also LaCour and Vavreck 2014).

Prior's (2013) second critique involves convergent validity, or the extent to which independent measures of the same concept produce similar results. He reported that 34 percent, 35 percent, and 35 percent of respondents checked off "Fox News" in waves 2, 4, and 5 of the 2008 NAES, whereas Nielsen data showed that only 10 percent and 13 percent of adults watched Fox News Channel (FNC) for at least 60 minutes during 2-week periods in April (wave 2) and October (wave 4) 2008. He noted that 26 percent, 30 percent, and 29 percent of adults watched at least 6 minutes of FNC per month during waves 2, 4, and 5, but dismissed those estimates as counting incidental viewers. Finally, he argued that the NAES measure failed to pick up increases in exposure late in the campaign.

In reply, Goldman et al. (2013) found that Nielsen has serious faults in its people-meter system and sampling (Milavsky 1992; Napoli 2003), and did not capture out-of-home viewing or exposure over the internet or on mobile devices (*Economist* 2013; Napoli 2003). Even if one does use Nielsen, the 6-minute cume for FNC corresponded to the NAES estimates. A rank ordering of programs by popularity also showed a strong parallel between the Nielsen and list-based estimates (Dilliplane et al. 2013). Finally, Goldman et al. (2013) reported a significant increase in the number of programs viewed as the 2008 election neared.

Prior's (2013) third critique concerns predictive validity, or whether a measure predicts future values of a criterion variable. He argued that political knowledge is an inappropriate criterion variable because news exposure is neither necessary nor sufficient to produce political learning and the true relationship between news exposure and political knowledge is unknown. He also suggested that Dilliplane et al. (2013) included weak controls for other media and omitted controls for exposure to political advertising, the party conventions, the presidential debates, and interpersonal discussion.

Goldman et al. (2013) responded that political television exposure does not need to be necessary or sufficient to have a causal impact on political knowledge and that the fixed effects panel models used by Dilliplane et al. (2013) provide unusually strong causal evidence (Allison 2009). The analyses included multiple controls for exposure to other media, and exposure to political television appropriately captured advertising, the conventions, and debates. Finally, interpersonal discussion is a potential mediator of media influence (Katz and Lazarsfeld 1955), yet controlling for it did not change the original findings (Goldman et al. 2013).

In his last critique, Prior (2013) argued that reliability is independent of, and less important than, validity, and that reliability estimates are upwardly biased due to correlated errors from overreporting. As noted by Goldman et al. (2013), however, Prior (2013) seems to confuse test-retest reliability, which is independent of validity, and true-score reliability, which represents the squared correlation between the observed score and true score, where "the distinction between reliability and validity does not exist" (Zaller 2002, 315). Moreover, even if correlated errors inflated the reliability estimates, this would not affect the predictive validity analyses because fixed effects regression *assumes* correlated errors (Allison 2009). With fixed effects, all stable individual differences, including consistent overreporting, automatically drop out.

Overall, the list-based measure has demonstrated strong reliability and validity and offers a clear improvement over traditional exposure measures, which have lower true-score reliability and have never been shown to predict within-person gains in political knowledge over time. Scholars have fruitfully employed the list-based approach to test a range of hypotheses, such as the effects of exposure to partisan (Dilliplane 2011, 2014; Moehler and Allen 2016), uncivil (Gervais 2014), and prejudice-reducing media (Goldman and Mutz 2014). The measure is also well-suited for linking individual-level survey data with content analyses, enabling more sophisticated tests of media influence (Jerit and Barabas 2011; Valkenburg and Peter 2013).

The List-Frequency Technique

As the list-based approach has grown in use, it has also become the subject of revision and extension. Most notably, Anderson et al. (2016) suggest that the measure could be improved in several ways, in particular by incorporating frequency of use. They dub the revised measure "the list-frequency technique," because it asks respondents how many days in the past week they used each source. The authors suggest that using the "past week" time frame minimizes the cognitive demands placed on respondents, increases observed variation over time in exposure, and improves precision in linkage analyses that combine self-reports with content analyses.

To assess the convergent validity of the list-frequency technique, Anderson et al. (2016) used an experiment embedded in a 2014 national survey in Denmark. Respondents were randomly assigned to the list-based measure or the list-frequency technique, with each including twelve newspapers, eight radio programs, eighteen television shows, and fifteen websites. The results showed no significant differences between conditions in the percent of respondents who used each media outlet, supporting the convergent validity of the list-frequency measure.

For tests of predictive validity, Anderson et al. (2016) used a two-wave panel survey that included list-frequency measures for three newspapers, four radio programs, six television shows, and five websites, with responses summed into indexes for each medium. LDVMs showed that the exposure measures significantly predicted greater wave 2 current events knowledge. To examine whether frequency of use added explanatory power, the authors created dichotomous variables for each exposure measure to approximate the list-based approach. Cross-sectional analyses showed that the list-frequency measures explained slightly more variance in knowledge, but only to a significant degree in the case of newspaper exposure.

Anderson et al.'s (2016) extension of the list-based approach is well-motivated and provides a rare empirical assessment of the value added of measuring frequency of media exposure. That said, based on the evidence provided, we do not yet see a clear advantage of adding frequency estimates to the list-based approach. It has long been a concern that respondents cannot accurately recall their frequency of media use; indeed, this was one of the main critiques of traditional exposure measures. People might be able to more reliably report their frequency of using individual sources than researcher-defined categories of news, but testing this idea requires true-score reliability estimates of the list-frequency measures using three or more waves of panel data. Such data would also enable tests of whether list-frequency measures predict individual-level knowledge *gains* over time. In sum, future research should continue to investigate the utility of the list-frequency technique.

Capturing Exposure to Political Websites

As people increasingly turn to the internet for political content, technological changes may alter our understanding of media effects theories (Mutz and Young 2011). Unfortunately, systematic assessment of measures of online media exposure remains largely neglected. In a rare exception, Guess (2015) tested the validity of three self-report measures of exposure to political websites. In three experiments, all fielded over Amazon.com's Mechanical Turk (MTurk), respondents answered one of three exposure measures. The first, a list-based measure, showed respondents a list of twenty-seven websites and asked them to check off those they had visited in the past thirty days for news. The second, a forced-choice measure, asked respondents to indicate "Yes" or "No" next to each site. The third measure provided respondents with a blank text box and asked them to list any websites they had visited for news in the past thirty days.

The tests of convergent validity relied on novel indicators of "true" exposure (Guess 2015). Two experiments used the "link classification technique," which takes advantage of web browsers automatically storing the links of sites that users visit, and then presenting visited and unvisited links in different colors. Respondents were shown a list of 155 hyperlinks, each labeled as "LINK" and asked to check off the visited sites shown in purple. As a robustness check, in the third experiment respondents were asked to install a widget that encoded their browsing history from the last 30 days. Based on these indicators, the list-based measure showed modest overreporting, as respondents reported visiting 2.33 more sites on average than the link-classification estimate. The forced-choice measure produced even more overreporting (4.15 more sites on average), while the open-ended measure did not produce significant overreporting (0.80 more sites on average) but did produce under-reporting of some sites. Overall, the open-ended and list-based measures did not differ in total misreporting.

The third experiment included tests of predictive validity based on the ability of each exposure measure to predict news recall, as measured by an index of the number of correct answers to three questions about recent news stories (Guess 2015). Controlling for general political knowledge and demographic characteristics, the list-based measure significantly predicted news recall, the open-ended measure produced an even larger effect, and the forced-choice measure had no impact. These findings support the predictive validity of the list-based and open-ended measures, but not the forced-choice measure.

For assessing exposure to political websites, the list-based and open-ended measures are promising. In both cases, we would recommend excluding reference to news in the questions given the ambiguous and debated meaning of the term. Although Guess (2015) found that the list-based measure showed some overreporting, the indictors of true exposure used for comparison likely underestimated exposure levels by capturing media use on only one device and only when respondents visited a website's homepage. Nonetheless, the list-based measure predicted news recall, and the open-ended measure performed even more strongly, despite small sample sizes and an outcome variable with limited variance. In terms of implementation in large-scale surveys, the list-based measure is the most feasible, though further research employing open-ended measures of media exposure is clearly warranted.

The Challenge of Measuring Social Media Exposure

Popular surveys suggest that a growing number of people get news at least some of the time from social media sites such as Facebook, Twitter, Snapchat, YouTube, Pinterest, Instagram, Tumblr, LinkedIn, and Reddit (e.g., Pew Research Center 2018). Although surveys increasingly include questions about social media use, research pertaining to the validity of these measures is sparse. To date, scholars have relied primarily on two approaches to gauging social media use. The first approach asks respondents whether they use social media at all. For example, Gottfried, Hardy, Holbert et al. (2017) asked respondents, "Do you ever use social networking sites such as Twitter or Facebook?" (yes or no). The second and more commonly used approach presents a list of social media sites and asks respondents how frequently they use each one to get news or political information, with response options ranging from the subjective (e.g., "never" to "all the time") to specific time periods (e.g., days per week).

The standard criterion for assessing the validity of media exposure measures is political knowledge gain, yet it is not clear that social media should be expected to produce learning of this kind. On the one hand, an experiment showed that social media exposure can produce gains in political knowledge (Bode 2016). On the other hand, research combining a survey with a Facebook application that collected participants' newsfeeds found no relationship between the presence of news stories on Facebook and individuals' knowledge of those same stories (Wells and Thorson 2017). Studies relying on self-reports offer mixed evidence as well. One cross-sectional analysis revealed a significant positive relationship between having ever used Twitter or Facebook and campaign knowledge (Gottfried et al. 2017), while another revealed a positive impact of using Twitter, but not Facebook, on knowledge of current events (Bode 2016). Moreover, five survey-based studies that asked respondents how often they followed news about politics or campaigns using each of several social media sites found no positive relationship between exposure to those sites and knowledge about politics or current events (Dimitrova et al. 2011; Gil de Zuniga, Weeks, and Ardevol-Abreu 2017; Groshek and Dimitrova 2011; Shehata and Stromback 2018; Wolfsfeld, Yarchi, and Samuel-Azran 2015). Taken together, prior research provides limited evidence that self-reported social media use is associated with political learning.

One interpretation of this weak relationship is that current measures of social media uselack predictive validity. Another interpretation is that knowledge gain is an inappropriate criterion variable for assessing the validity of these measures. Many suggest that social media use is better suited for promoting political participation, as indicated by a large-scale Facebook experiment, however it is not clear if naturally occurring social media use increases participation(Bond et al. 2012). Although cross-sectional correlations between self-reported social media use and political participation abound, panel studies that provide stronger causal evidence are less common and reveal fewer significant associations (Boulianne 2015).

The inconsistency with which self-reported social media use predicts gains in either political knowledge or participation suggests that measurement error may be at least partly to blame (Bartels 1993). Asking respondents if they use social media for getting news introduces error due to widely varying interpretations of what counts as news. And asking about frequency of use introduces error owing to imperfect memory, especially for common behaviors like social media use. In addition, many measures only ask whether respondents purposefully use social media for getting news or following politics, which may miss a potentially large number of people for whom exposure to politics is incidental. A list-based measure which asks respondents whether they have heard anything about an electoral campaign or politics on each of several social media sites could reduce measurement error while still capturing incidental exposure. Unfortunately, a list-based measure cannot distinguish between exposure to different messages and types of content within a given social media site. For this purpose, self-reports combined with passive tracking data of each respondent's social media feed will likely be necessary to test theories of individual-level media effects.

Research Design

The question of how to measure political media exposure is interrelated with how to capture the effects of exposure on attitudes and behaviors. An ideal research design maximizes observed variation in the independent and dependent variables, while minimizing threats to causal inference (i.e., spuriousness and reverse causality). For example, the challenge of observing variation in media exposure is particularly notable in the debate over the influence of US presidential campaigns on vote choice. While the public perceives campaigns as powerful, academic research has shown small-to-null effects (Mutz 2012). As Zaller (1996) argued, however, there could be massive effects that are difficult to observe. Because presidential campaigns tend to be evenly matched, their effects may cancel out, leaving little observed variation in net exposure to campaign appeals. This problem is compounded by the static design of most election studies, which, combined with small samples and low statistical power (Zaller 2002), severely limits scholars' ability to capture the effects of media exposure on vote choice.

The use of cross-sectional surveys in most observational media effects studies also undercuts the ability of scholars to make causal inferences due to numerous threats to internal validity (Banducci et al. 2017). Although comparing individuals with higher levels of media exposure to those with lower levels of exposure exploits substantial between-person variation, any differences in attitudes or behaviors between those individuals may be due to any number of confounding factors that may produce spurious associations (Bartels 1993). Importantly, no amount of statistical wizardry can solve this fundamental problem of causal inference with between-person comparisons. Even sophisticated matching techniques, which rely on fewer modeling assumptions than standard regression analyses, cannot control for unmeasured and/or unobservable factors, not to mention reverse causality (Banducci et al. 2017).

A stronger design that has come into greater use is the *rolling* cross-sectional survey (Brady and Johnston 2006), which randomly assigns respondents a date of interview over a period of weeks or months. Assuming large enough daily samples, researchers can identify the effects of campaign events and abrupt shifts in media coverage, so long as other environmental influences are ruled out (Johnston, Hagen, and Jamieson 2004). Analyses of campaign dynamics take place at the aggregate-level, usually through graphical analysis of daily trends in beliefs and vote intentions alongside trends in media coverage (Kenski, Hardy, and Jamieson 2010). The major limitation of these analyses is the lack of direct evidence connecting media exposure with changes in attitudes at the individual level. Although researchers can compare trends among those with differing levels of self-reported media exposure, such between-person comparisons again raise the specter of spuriousness due to individual differences.

Ultimately, the best strategy for ruling out alternative explanations stemming from individual differences is to discard potentially contaminated between-person variance and instead rely strictly on within-person variance. For this task, panel data are necessary but not sufficient. The traditional approach to analyzing panel data using lagged dependent variable models (LDVMs) still relies entirely on between-person variance. Researchers often describe LDVMs as demonstrating individual-level change, but this is a misnomer as these models only reveal change in the rank-order of individuals. As a result, LDVMs often provide results that make little substantive sense when compared to the underlying distribution of responses (Allison 1990). By contrast, fixed effects regression models rely solely on within-person variance and assess whether each individual has changed in her attitudes over time (Allison 2009).

The benefits of fixed effects regression for disentangling the impact of media exposure from individual differences are enormous: by using each respondent as her own control, the stable effects of *all* individual differences automatically drop out. Each respondent is compared to herself at an earlier point in time, so the main effects of factors like education, political interest, gender, and party identification are all controlled. Indeed, if one tries to include stable variables in a fixed effects model they will not appear in the output. More important than controlling for factors that we already know about and can measure is that fixed effects models automatically control for the stable effects of every other factor that we do not know about and/or cannot measure (i.e., unobserved heterogeneity). Including a variable for panel wave efficiently captures the sum total of all other changes (i.e., period effects) (Halaby 2004).

The remaining routes through which spuriousness can influence fixed effects models are limited. Because the main effects of individual differences automatically drop out, only their time-varying components can be influential. These time-varying components can be modeled with interactions between individual difference variables and survey wave. Unsurprisingly, including these controls rarely alters the size or significance of media exposure variables (Dilliplane 2011; Dilliplane et al. 2013; Goldman 2012; Goldman and Mutz 2014). Factors that change over time represent another plausible, albeit unlikely, influence. With the wave variable already capturing factors that change uniformly over time, a spurious confounder would have to change differentially among individuals, and those changes would have to be correlated with changes in both the independent and dependent variables. Few factors plausibly fit this bill, and including them typically leaves most estimates of media impact unaffected.

With fixed effects, one can model the influence of media exposure in two ways. The first is a change-on-change model that requires repeated measurements of media exposure. For example, using three waves of panel data Dilliplane et al. (2013) found that within-person change in the number of political TV programs viewed significantly predicted within-person change in knowledge of candidate issue positions during the 2008 US presidential campaign. The second type of model examines the time-varying effects of a stable indicator of media exposure. Here, Dilliplane et al. (2013) created a scale that averaged self-reports of media exposure across all three waves and then examined the interaction of the exposure scale with the wave variable. They found a positive interaction, indicating that stable levels of exposure predicted within-person gains in knowledge. Each modeling strategy has distinct benefits. The change-on-change model does a superior job of controlling for unobservables, though limited within-person variation in exposure can produce conservative estimates of effect size (Allison 2009). Using stable exposure, on the other hand, captures the habitual character of media exposure and can help rule out reverse causality.

Given the range of research designs available, some suggest using as many as possible—including between-subjects models—in order to demonstrate the robustness of one's results (Banducci et al. 2017). Yet given that the internal validity of between-subjects models is widely suspect, it is not evident why the results would be informative. All research designs are not equally able to rule out alternative explanations (Campbell and Stanley 1963), which suggests that scholars should use the best possible research design available for the purpose at hand. For example, Goldman and Mutz (2014) used the 2008 NAES internet panel survey, which combined the panel and rolling cross-sectional designs, to link respondents' date of interview with the results of a content analysis. A fixed effects model of within-person change then showed that increases over time in political TV viewing produced the largest declines in racial prejudice when accompanied by increases over time in coverage of Barack Obama refuting anti-Black stereotypes. Although one could also carry out between-subjects analyses of these data, those results—whatever they showed—would provide weak causal evidence.

Conclusion

Traditional measures of media exposure have long been debated and widely ridiculed on the grounds of low reliability and validity. Especially since Price and Zaller's (1993) landmark study, the ability of traditional frequency measures to capture the extent to which people watch, read, or listen to the news has been suspect, leading Bartels (1993, 267) to famously proclaim that, "The state of research on media effects is one of the most notable embarrassments of modern social science." Perhaps this was just the wake-up call scholars needed. As this review shows, although consensus remains elusive, there has been substantial progress in the measurement and evaluation of the effects of media exposure.

With regard to measurement, early critiques of traditional measures are less clear-cut than many assumed. Findings widely interpreted as showing low reliability of traditional measures fare well relative to many survey measures of political constructs (Bartels 1993). Evidence of low predictive validity (Price and Zaller 1993) has also been interpreted by some as showing sensible, if modest, relationships (Althaus and Tewksbury 2007a, 2007b). However, no study to our knowledge has used panel data to examine whether change over time in the frequency of news exposure predicts individual knowledge gain. But perhaps most surprising is recent evidence of convergent validity: despite some overreporting, an hours-per-day measure of TV news exposure still distinguished between those with high versus low levels of actual news viewing based on passive tracking data (LaCour and Vavreck 2014).

Nonetheless, the utility of traditional exposure measures has declined drastically with the end of the broadcast era of TV news (Delli Carpini and Williams 2011). The menu of media options has grown exponentially, the line between news and entertainment has blurred, and audiences have fragmented. Scholars now want to understand the effects of exposure to the varied forms of politically relevant content available across platforms, devices, and viewing behaviors. For these purposes, list-based measures that are able to capture exposure to the specific TV programs, newspapers, radio shows, and websites people use are far more useful than indicators of exposure to “national TV news” or “news on the internet.” List-based measures have shown strong reliability and validity, and efforts to improve and expand on this approach have led to applications across a variety of media contexts.

Recent advances notwithstanding, we see more need than ever for a renewed scholarly focus on issues of measurement and research design, which still receive far less attention than they deserve. For instance, scholars have only just begun to investigate the complexity of citizens’ media diets (Moehler and Allen 2016), consumption patterns across convergent media on mobile devices (Ohme, Albaek, and de Vreese 2016), and multitasking behaviors (Gottfried et al. 2017). Yet despite these challenges, understanding how people use and respond to political media is necessary in order to confront many of the most important issues facing democratic institutions. Although the debate over how to measure media exposure will undoubtedly continue, it is just as clear that scholars have never been better positioned to make use of this key scientific concept to promote societal well-being.

References

Allison, Paul D. 1990. “Change Scores as Dependent Variables in Regression Analysis.” In *Sociological Methodology*, edited by Clifford Clogg, 93–114. Oxford: Basil Blackwell.

Allison, Paul D. 2009. *Fixed Effects Regression Models*. Thousand Oaks, CA: Sage Publications.

Althaus, Scott L., and David H. Tewksbury. 2007a. “Do We Still Need Media Use Measures at All?” In *Improving Public Opinion Surveys Book: Interdisciplinary Innovation and the*

American National Election Studies, edited by John H. Aldrich and Kathleen M. McGraw, 158–174. Princeton, NJ: Princeton University Press.

Althaus, Scott L., and David H. Tewksbury. 2007b. “Toward a New Generation of Media Use Measures for the ANES: Report to the Board of Overseers American National Election Studies.” *American National Election Studies Pilot Study Report*.

Anderson, Kim, Claes H. de Vreese, and Erik Albaek. 2016. “Measuring Media Diet in a High-Choice Environment: Testing the List-Frequency Technique.” *Communication Methods and Measures* 10, nos. 2–3: 81–98.

Banducci, Susan, Martijn Schoonvelde, Daniel Stevens, Jason Barabas, Jennifer Jerit, and William Pollock. 2017. “Model Selection in Observational Media Effects Research: A Systematic Review and Validation of Effects.” *Political Science* 69, no. 3: 227–246.

Bartels, Larry M. “Messages Received: The Political Impact of Media Exposure.” 1993. *American Political Science Review* 87, no. 2: 267–285.

Bode, Leticia. 2016. “Political News in the News Feed: Learning Politics from Social Media.” *Mass Communication and Society* 19, no. 1: 24–48.

Bond, Robert M., Christopher J. Fariss, Jason J. Jones, Adam D. I. Kramer, Cameron Marlow, Jaime E. Settle, et al. 2012. “A 61-Million-Person Experiment in Social Influence and Political Mobilization.” *Nature* 489, no. 7415: 295–298.

Boulianne, Shelley. 2015. “Social Media Use and Participation: A Meta-Analysis of Current Research.” *Information, Communication & Society* 18, no. 5: 524–538.

Brady, Henry E., and Richard Johnston. 2006. “The Rolling Cross-Section and Causal Attribution.” In *Capturing Campaign Effects*, edited by Henry E. Brady and Richard Johnston, 164–195. Ann Arbor: University of Michigan Press.

Campbell, Donald T., and Julian C. Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Belmont, CA: Wadsworth Publishing.

Chaffee, Steven H., and Joan Schleuder. 1986. "Measurement and Effects of Attention to Media News." *Human Communication Research* 13, no. 1: 76–107.

Chang, LinChiat, and Jon A. Krosnick. 2003. "Measuring the Frequency of Regular Behaviors: Comparing the "Typical Week" to the "Past Week." *Sociological Methodology* 33: 55–80.

"Counting Couch Potatoes." February 9, 2013. *The Economist*. https://www.economist.com/business/2013/02/09/counting-couch-potatoes.

De Vreese, Claes H., and Peter Neijens. 2016. "Measuring Media Exposure in a Change Communications Environment." *Communication Methods and Measures* 10, nos. 2–3: 69–80.

Delli Carpini, Michael X., and Bruce A. Williams. 2011. *After Broadcast News: Media Regimes, Democracy, and the New Information Environment*. Cambridge, UK: Cambridge University Press.

Dilliplane, Susanna. 2011. "All the News You Want to Hear: The Impact of Partisan News Exposure on Political Participation." *Public Opinion Quarterly* 75, no. 2: 287–316.

Dilliplane, Susanna. 2014. "Activation, Conversion, or Reinforcement? The Impact of Partisan News Exposure on Vote Choice." *American Journal of Political Science* 58, no. 1: 79–94.

Dilliplane, Susanna, Seth K. Goldman, and Diana C. Mutz. 2013. "Televised Exposure to Politics: New Measures for a Fragmented Media Environment." *American Journal of Political Science* 57, no. 1: 236–248.

Dimitrova, Daniela V., Adam Shehata, Jesper Stromback, and Lars W. Nord. 2011. “The Effects of Digital Media on Political Knowledge and Participation in Election Campaigns: Evidence from Panel Data.” *Communication Research* 41, no. 1: 95–118.

Edgerly, Stephanie, Emily K. Vraga, Leticia Bode, Kjerstin Thorson, and Esther Thorson. 2017. “New Media, New Relationship to Participation? A Closer Look at Youth News Repertoires and Political Participation.” *Journalism & Mass Communication Quarterly* 95, no. 1: 192–212.

Eveland, William P., Andrew F. Hayes, Dhavan V. Shah, and Nojin Kwak. 2005. “Understanding the Relationship between Communication and Political Knowledge: A Model Comparison Approach Using Panel Data.” *Political Communication* 22, no. 4: 423–446.

Eveland, William P., Myiah J. Hutchens, and Fei Shen. 2009. “Exposure, Attention, or ‘Use’ of News? Assessing Aspects of the Reliability and Validity of a Central Concept in Political Communication Research.” *Communication Methods and Measures* 3, no. 4: 223–244.

Gervais, Bryan T. 2014. “Following the News? Reception of Uncivil Partisan Media and the Use of Incivility in Political Expression.” *Political Communication* 31, no. 4: 564–583.

Gil de Zuniga, Homero, Brian Weeks, and Alberto Ardevol-Abreu. 2017. “Effects of the News-Finds-Me Perception in Communication: Social Media Use Implications for News Seeking and Learning about Politics.” *Journal of Computer-Mediated Communication* 22: 105–123.

Goldman, Seth K. 2012. “Effects of the 2008 Presidential Campaign on White Racial Prejudice.” *Public Opinion Quarterly* 76, no. 4: 663–687.

Goldman, Seth K., and Diana C. Mutz. 2014. *The Obama Effect: How the 2008 Campaign Changed White Racial Attitudes*. New York: Russell Sage Foundation.

Goldman, Seth K., Diana C. Mutz, and Susanna Dilliplane. 2013. “All Virtue is Relative: A Response to Prior.” *Political Communication* 30, no. 4: 635–653.

Gottfried, Jeffrey A., Bruce W. Hardy, R. Lance Holbert, Kenneth M. Winneg, and Kathleen Hall Jamieson. 2017. “The Changing Nature of Political Debate Consumption: Social Media, Multitasking, and Knowledge Acquisition.” *Political Communication* 34, no. 2: 172–199.

Groshek, Jacob, and Daniela Dimitrova. 2011. “A Cross-Section of Voter Learning, Campaign Interest and Intention to Vote in the 2008 American Election: Did Web 2.0 Matter?” *Communication Studies Journal* 9, no. 1: 355–375.

Guess, Andrew M. 2015. “Measure for Measure: An Experimental Test of Online Political Media Exposure.” *Political Analysis* 23, no. 1: 59–75.

Halaby, Charles N. 2004. “Panel Models in Sociological Research: Theory into Practice.” *Annual Review of Sociology* 30: 507–544.

Heise, David R. 1969. “Separating Reliability and Stability in Test-Retest Correlation.” *American Sociological Review* 34, no. 1: 93–101.

Huesmann, L. Rowell, Jessica Moise-Titus, Cheryl-Lynn Podolski, and Leonard Dr. Eron. 2003. “Longitudinal Relationships Between Children’s Exposure to TV Violence and their Aggressive and Violent Behavior in Young Adulthood: 1977–92.” *Developmental Psychology* 39, no. 2: 201–221.

Jerit, Jennifer, and Jason Barabas. 2011. “Exposure Measures and Content Analysis in Media Effects Studies.” In *The Oxford Handbook of American Public Opinion and the Media*, edited by Robert Y. Shapiro and Lawrence R. Jacobs, 139–155. New York: Oxford University Press.

Johnston, Richard, Michael G. Hagen, and Kathleen Hall Jamieson. 2004. *The 2000 Presidential Election and the Foundations of Party Politics*. Cambridge, UK: Cambridge University Press.

Katz, Elihu, and Paul F. Lazarsfeld. 1955. *Personal Influence: The Part Played by People in the Flow of Mass Communications*. New York: Free Press.

Kenski, Kate, Bruce W. Hardy, and Kathleen Hall Jamieson. 2010. *The Obama Victory: How Media, Money, and Message Shaped the 2008 Election*. Oxford: Oxford University Press.

Kruikemeier, Sanne, and Adam Shehata. 2017. "News Media Use and Political Engagement among Adolescents: An Analysis of Virtuous Circles Using Panel Data." *Political Communication* 34, no. 2: 221–242.

LaCour, Michael J., and Lynn Vavreck. 2014. "Improving Media Measurement: Evidence from the Field." *Political Communication* 31, no. 3: 408–420.

Milavsky, J. Ronald. 1992. "How Good is the A. C. Nielsen People-Meter System? A Review of the Report by the Committee on Nationwide Television Audience Measurement." *Public Opinion Quarterly* 56: 102–115.

Moehler, Devra, and Douglas Michael Allen. 2016. "The Media Diet Imbalance Score: A Measure of Aggregate Media Diet." *Communication Methods and Measures* 10, no. 1: 4–12.

Mutz, Diana C. 2012. "The Great Divide: Campaign Media in the American Mind." *Daedalus* 141, no. 4: 83–97.

Mutz, Diana C., and Lori Young. 2011. "Communication and Public Opinion: Plus Ca Change?" *Public Opinion Quarterly* 75, no. 5: 1018–1044.

Nagler, Rebekah H. 2018. "Measurement of Media Exposure." In *International Encyclopedia of Communication Research Methods*, edited by Jorg Matthes. Hoboken, NJ: Wiley-Blackwell and the International Communication Association.

Napoli, Philip M. 2003. *Audience Economics: Media Institutions and the Audience Marketplace*. New York: Columbia University Press.

Ohme, Jakob, Erik Albaek, and Claes H. de Vreese. 2016. "Exposure Research Going Mobile: A Smartphone-Based Measurement of Media Exposure to Political Information in a Convergent Media Environment." *Communication Methods and Measures* 10, nos. 2–3: 135–148.

Pew Research Center. 2018. "News Use Across Social Media Platforms 2018." September 10. http://www.journalism.org/2018/09/10/news-use-across-social-media-platforms-2018/. Retrieved October 9, 2018.

Price, Vincent, and John Zaller. 1993. "Who Gets the News? Alternative Measures of News Reception and Their Implications for Research." *Public Opinion Quarterly* 57, no. 2: 133–164.

Prior, Markus. 2009a. "Improving Media Effects Research through Better Measurement of News Exposure." *The Journal of Politics* 71, no. 3: 893–908.

Prior, Markus. 2009b. "The Immensely Inflated News Audience: Assessing Bias in Self-Reported News Exposure." *Public Opinion Quarterly* 73, no. 1: 130–143.

Prior, Markus. 2013. "The Challenge of Measuring Media Exposure: Reply to Dilliplane, Goldman, and Mutz." *Political Communication* 30, no. 4: 620–634.

Schwarz, Norbert, and Daphna Oyserman. 2001. "Asking Questions about Behavior: Cognition, Communication, and Questionnaire Construction." *American Journal of Evaluation* 22, no. 2: 127–160.

Shehata, Adam, and Jesper Stromback. 2018. "Learning Political News from Social Media: Network Media Logic and Current Affairs News Learning in a High-Choice Media Environment." *Communication Research*.

Slater, Michael D. 2004. "Operationalizing and Analyzing Exposure: The Foundation of Media Effects Research." *Journalism & Mass Communication Quarterly* 81, no. 1: 168–183.

Tewksbury, David, Scott L. Althaus, and Matthew V. Hibbing. 2011. "Estimating Self-Reported News Exposure Across and Within Typical Days: Should Surveys Use More Refined Measures." *Communication Methods and Measures* 5, no. 4: 311–328.

Valkenburg, Patti M, and Jochen Peter. 2013. "Five Challenges for the Future of Media-Effects Research." *International Journal of Communication* 7: 197–215.

Wells, Chris, and Kjerstin Thorson. 2017. "Combining Big Data and Survey Techniques to Model Effects of Political Content Flows in Facebook." *Social Science Computer Review* 35, no. 1: 33–52.

Wiley, David E., and James A. Wiley. 1970. "The Estimation of Measurement Error in Panel Data." *American Sociological Review* 35, no. 1: 112–117.

Wolfsfeld, Gadi, Moran Yarchi, and Tal Samuel-Azran. 2015. "Political Information Repertoires and Political Participation." *New Media & Society* 18, no. 9: 2096–2115.

Wonneberger, Anke, Klaus Schoenbach, and Lex van Meurs. 2013. "Dimensionality of TV-News Exposure: Mapping News Viewing Behavior with People-Meter Data." *International Journal of Public Opinion Research* 25, no. 1: 87–107.

Zaller, John. 1996. “The Myth of Massive Media Impact Revived: New Support for a Discredited Idea.” In *Political Persuasion and Attitude Change*, edited by Diana C. Mutz, Richard A. Brody, and Paul M. Sniderman, 17–79. Ann Arbor: University of Michigan Press.

Zaller, John. 2002. “The Statistical Power of Election Studies to Detect Media Exposure Effects in Political Campaigns.” *Electoral Studies* 21, no. 2: 297–329.